INPUT NEEDS FOR DOWNSCALING OF CLIMATE DATA

DISCUSSION PAPER

Prepared For:
California Energy Commission
Public Interest Energy Research Program

Prepared By:
Tom M. L. Wigley
Consultant

January 2004

Prepared For:
California Energy Commission

Guido Franco,
Project Manager

Kelly Birkinshaw,
Program Area Lead

Terry Surles,
PIER Program Manager

Robert L. Therkelsen,
Executive Director

Prepared By:
Tom M. L. Wigley*
Consultant
Louisville, CO 80027-1064

*Tom Wigley is a Senior Scientist at the National Center for Atmospheric Research (NCAR), Boulder, Colorado

Acknowledgments

Contents

List of Figures

Abstract

This report addresses the external data-input needs for climate data downscaling, with emphasis on the U.S. western region and a focus on producing high-spatial-resolution climate data for the future. Three downscaling methods are considered: two types of dynamical downscaling, using regional climate models (RCMs) or high-resolution global atmospheric General Circulation Models (AGCMs), and statistical downscaling. In all cases, the quality of the downscaled product must be dependent on the quality of the driver Atmosphere-Ocean GCM (AOGCM). One must therefore validate both the downscaling method and the driver model. RCMs can be validated using reanalysis data for the lateral and surface boundary conditions, AGCMs can be validated by driving them with observed surface boundary conditions, and statistical models can be validated using data that were not employed in model calibration. AOGCM validation is performed by comparing their simulations of present-day climate with observed and reanalysis data. Emphasis here should be placed on the driver GCM's ability to simulate reliably present-day conditions for both the boundary conditions and fluxes and the natural modes of variability (such as ENSO) that are important in the study region. Models that perform less well in present-day validation may still be useful for investigating changes in climate. The choice of driver GCM should consider not only present-day validation performance, but also the model's simulations of future changes in climate over the study region. With individual models there are problems in identifying anthropogenic precipitation signals because of the large magnitude of natural variability that is common to the U.S. western region. These signal-to-noise ratio problems are exacerbated by inter-model differences in their responses to external forcing. Because different models give a wide range of results for future climate, particularly for precipitation, models should be chosen to span the range of future possibilities.

1.0   INTRODUCTION

Global coupled Atmosphere-Ocean General Circulation Models (AOGCMs) currently used for projecting future climate have a grid box size of 100–200 km. Many of these models are able to simulate present-day climate well on spatial scales of 1000 km upwards, and the best models provide reasonable representations of the climate on somewhat smaller scales. Their grid-box resolution, however, cannot capture the details of orography nor resolve important cyclonic disturbances or similar-sized circulation features. This precludes an accurate representation of the climate on scales of individual grid boxes. For many impacts models, however, information is required on sub-grid scales of 10 to 100km (referred to here as the *local* to *regional* scale). The method for producing local-to-regional scale information from larger-scale GCM data is called *downscaling.*

Two downscaling methods are commonly used, dynamical and statistical downscaling. *Dynamical downscaling* methods include: the use of a limited-area, high-resolution Regional Climate Model (RCM) nested within and driven by time-dependent lateral and lower boundary conditions from a global AOGCM; the use of a global model with variable spatial resolution (a stretched-grid atmospheric GCM, or SG-AGCM); or the use of a high-resolution global AGCM in time-slice experiments driven by AOGCM forcings and surface boundary conditions. It is also possible to use a hybrid approach, in which an ultra-high-resolution RCM is driven by lateral boundary conditions from a high-resolution global AGCM, which in turn in is forced by lower boundary conditions from a coarse-resolution AOGCM. There are two types of RCM-based downscaling, depending on whether the RCM results feed back to the driver GCM (two-way nesting) or not (one-way nesting). *Statistical downscaling* involves the derivation, validation, and application of a statistical model (usually based on regression analysis) that relates regional-scale predictors to local-scale climate variables.

2.0   DYNAMICAL DOWNSCALING

This section discusses the input needs for dynamical downscaling. In the case of RCM-based downscaling, this means the production of high-resolution climate data through the use of a limited-area, high-resolution RCM driven by time-dependent lateral and lower boundary conditions from either reanalysis data (for simulations of present-day climate) or a global-scale, coarse-resolution AOGCM (for simulations of both present-day or future climate). In the case of downscaling with a high-resolution global AGCM, only lower boundary conditions are needed.

When used for downscaling, RCMs require, as inputs, specifications of the following parameters:

- the initial state of the atmosphere and surface
- time variations of any external forcing agents
- the time-varying lateral fluxes of heat, mass, moisture, momentum and forcing agents (such as aerosols) derived from a global AOGCM or reanalysis
- time variations of all surface variables that are not produced prognostically by the RCM (where *prognostically* means 'derived internally by the model's physics and chemistry')
- time variations in all atmospheric composition changes that are not produced prognostically by the RCM.

Downscaling using high-resolution global AGCMs requires all of the above inputs, except for lateral fluxes.

These requirements are discussed below under four headings: Synoptic Climatology, Existing Studies, Driver GCMs, and Aerosols.

2.1.   Synoptic Climatology

The climate of the western United States is determined by the large-scale influences of the Aleutian Low and the north Pacific subtropical high, which control the prevailing winds and storm tracks and their seasonal changes (Cayan and Roads 1984; Mitchell and Blier 1997). These aspects are modified in major ways by the complex geography and orography of the region. Orographic influences produce enhanced precipitation on the western sides of the coastal ranges, the Cascades and the Sierra Nevada, and pronounced rain shadows on the eastern (lee) sides. Strong precipitation gradients occur over distances of less than 100km—much less than the resolution of current global AOGCMs.

For precipitation, seasonal variability is controlled by the sources of moisture. On western slopes, the main moisture source is from the Pacific. This source affects the region primarily in the winter months. Northward movement of the Aleutian Low causes a fairly abrupt reduction in precipitation as the seasons shift from winter to spring to summer. On eastern slopes, the main moisture sources are transport at upper levels from the Gulf of Mexico and lower-level moisture transport from the lower latitudes of the eastern Pacific and the Gulf of California associated with the development of the North American monsoon in June and July (see, e.g.,

Chen et al. 1999). Inter-annual variability is strongly influence by the El Niño/Southern Oscillation (ENSO) (see, e.g., Cayan 1996; Trenberth and Guillemot 1996; Cayan et al. 1999). ENSO influences are manifest by changes in the intensity of the Aleutian Low and the subtropical high (see, e.g., Trenberth and Hurrell 1994), a pattern of changes that is commonly called the Pacific-North America (PNA) pattern. Longer time scale changes appear to be associated with the Pacific Decadal Oscillation (PDO) (Mantua and Hare 2002).

2.2.   Existing Studies

The first application of an RCM to produce small-scale climate information is the work of Dickinson et al. (1989) and Giorgi and Bates (1989). Many improvements have been made since then, and the current state of the art has been reviewed in the IPCC Third Assessment Report (TAR) (Giorgi et al. 2001) and by Leung et al. (2003a). An extensive intercomparison project has been carried out using a North American domain (Program to Intercompare Regional Climate Simulations, or PIRCS) (Takle et al. 1999), and a similar project is under way in Europe (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects, or PRUDENCE. See http://prudence.dmi.dk). These studies provide extensive information on currently available RCMs. Recent work focusing on the western United States (e.g., Giorgi et al. 1993a; Kim 1997; Chen et al. 1999; Leung and Ghan 1999a,b; Kim et al. 2000; Denis et al. 2002, 2003; and, particularly, Leung et al. 2003b,c) has shown that RCMs are able to reproduce many characteristics of the region's present-day local climate (see above) with remarkable skill.

Downscaling via high-resolution global AGCMs is a less mature and thus less well-documented approach. As discussed in Section 4 below, Duffy et al. (2003) performed and evaluated simulations of the present climate with the CCM3 global AGCM at resolutions of ~75 km and ~50 km. Govindasamy et al. (2003) downscaled a $2xCO_2$ climate using the CCM3 model at T170 truncation (~75 km resolution) driven by lower boundary conditions from the coarse-resolution AOGCM known as the Parallel Climate Model (PCM).

2.3.   Driver GCMs

A number of factors need to be considered in choosing a driver GCM to specify lateral and surface boundary conditions and free-atmosphere composition changes: availability of suitable experiments; availability of data with suitable temporal resolution; quality of the GCM; and parameterization bias.

2.3.1.   Availability of suitable experiments

In order to study climate <u>change</u> over the region of interest (California specifically, but, more generally, the western one-third of the United States, where climatic and hydrological conditions may significantly affect the climate of California), one must have input data from two simulations for the present (one for validation purposes and one as the baseline for future changes) and one for some future period spanning a minimum of 10 years.

For validation purposes, present-day data should come from reanalyses. Reanalyses are essentially syntheses of observational data from all readily available sources. To ensure internal consistency between different data types and sources, to fill in any data gaps, and to produce a high-resolution gridded data set, the observed data are combined with weather forecast information in a statistically optimized way. They provide the best possible observed boundary data set for driving an RCM if we seek to see how well the RCM can reproduce the details of observed present-day climate within its spatial domain. Boundary conditions from reanalyses may be assumed to be of high reliability, so the RCM's simulations of the magnitudes and patterns of climate within the RCM domain, when compared with observations, provide a crucial test of the model. Furthermore, use of a standard set of reanalysis boundary data for evaluating a set of RCMs can provide a means of deciding which RCMs might be preferred for predicting the details of future climate, since it is often assumed that present-day simulation skill is a pre-requisite for skill in future projections. It should be noted, however, that present-day skill in a RCM is not a sufficient condition for skill in future projections, since this is also dependent on the credibility of the driver AOGCM.

Until recently, either the NCEP/NCAR reanalysis (Kalnay et al. 1996) or the ECMWF ERA-15 reanalysis (Gibson et al. 1997) has been used for this purpose. A new and improved reanalysis product, ERA-40 (Simmons and Gibson 2003), has just become available, and any new RCM intercomparison exercise should consider using these data. ERA-40 is demonstrably superior to previous reanalysis products (see, e.g., Santer et al. 2004).

Future simulations must almost certainly use state-of-the-art coupled AOGCMs to define the RCM boundary conditions, because only fully coupled models can capture the range of internal variability (such as that related to ENSO) that is characteristic of California's climate. To obtain information about future climate, one could employ two methods: (1) use RCM results driven by a future AOGCM simulation directly, or (2) add the simulated change to the observed present-day (e.g., reanalysis) climate state. In the latter case, the change must be defined as the difference between RCM results driven by a future AOGCM simulation and those driven by a present-day simulation. Both methods have disadvantages.

In the first method, a minimum condition must be that the AOGCM simulation of present-day climate be accurate. If the AOGCM's present is biased, this will almost certainly lead to a bias in its future state. An important aspect of model (AOGCM) testing, therefore, is to compare RCM results driven by a reanalysis with those driven by the AOGCM's "present" (a procedure that applies equally to statistical downscaling).

Testing the ability of the AOGCM/RCM to simulate present-day climate requires, at least, a comparison of the AOGCM present-day simulation with observational and/or reanalysis data. It should be noted that this test may be unfair for the AOGCM, because most AOGCM simulations of the present do not include all of the present-day forcing factors (some of which are highly uncertain). As an example, carbonaceous aerosols (i.e., *soot*) are thought to be an important regional-scale forcing factor, but the time-varying distribution of these aerosols (which, ideally, should be an input data set for the AOGCM) is not well known. Even if it were,

the ability of AOGCMs to simulate the effects of these aerosols on weather and climate reliably has yet to be established.

The second method assumes that climate changes can be reliably simulated even if there are errors in the initial state. It is only possible to use monthly or longer time scale data to define the change here, since differencing on shorter time scales would lead to unrealistic high-frequency variability. In other words, for simulations of weather timescale variability, as required by many impacts models, it is not possible to simulate the future by adding modeled weather-time-scale changes to the present state. If only monthly time scale changes can be used, the future climate state would necessarily echo the baseline (present-day) state in terms of high-frequency (e.g., daily) variability.

Changes in variability could be deduced by comparing present and future RCM simulations, but it is not clear how these changes could be incorporated into the simulated future state. This problem may or may not be serious, however, since AOGCM changes in variability are frequently below the threshold for statistical significance and because such changes vary markedly from model to model (see, e.g., Wilby and Wigley 2002) and so must be treated more circumspectly than changes in the mean state. A more serious problem with this second method is the possibility that the driver AOGCM may show significant drift in the spatial patterns of change (see discussion later in this section).

For defining climate change on the monthly time scale there are two options. The first is to use idealized experiments, such as those employed in the Coupled Model Intercomparison (CMIP) project (Covey et al. 2003). Here, for some 20 models, monthly data are available for a standard (unforced) control run and a parallel run in which CO_2 concentration was increased at a compound rate of 1% per year (corresponding to a linear increase in radiative forcing). Note that these runs have no particular calendar dates associated with them. It is common practice with these data to define the change in climate by comparing RCM results driven by years 61–80 from the control run with years 61–80 from the 1% CO_2 run (where both runs begin in year zero with the same initial state). To apply these results to the future, a scaling assumption may be used—that is, if the results correspond to a certain global-mean warming amount, then they may be scaled up or down to simulate other warming magnitudes (Santer et al. 1990).

Implicit in this simple definition of climate change are three important considerations. First is the choice of 20 years for the simulations. This choice involves a compromise. It is important to use a period sufficiently long to give a reasonably stable estimate of the climate statistics (such as the mean climate and its inter-annual variability, and characteristics like the frequency of wet days and the distribution of wet-day amounts) for both the initial and final states; but it is also important with time-dependent forcing to have a period short enough to minimize trends over the analysis interval. The second consideration is the time period between the initial and final states (effectively 70 years here), which defines the magnitude of the climate-change signal. It is important to have a signal that is as large as possible, relative to the noise of natural variability (i.e., to have a large signal-to-noise ratio, or SNR). In California, precipitation variability is substantial, so even with a strong signal the SNR may not be large, and one may not be able to

see a significant change over short time periods. This issue is discussed further in Appendix 1. The <u>third</u> consideration is the use of the same time interval (years 61–80) for the two cases, because many AOGCMs show significant drift at sub-global scales in their control runs. If the drift is common to both the control and perturbation experiments (as confirmed by analyses with the NCAR/DOE Parallel Climate Model—at least for this particular model), then any error associated with drift will cancel out.

The second (and preferred) option is to use results from experiments with more realistic external forcing, such as simulations following the IPCC SRES emissions scenarios. The above signal-to-noise and drift issues apply equally to these cases. The possible effects of drift are usually ignored here, and a standard procedure is to compare a "present" period (such as 1980–1999) with a future period (such as 2080–2099) in an experiment that begins at some nominal "pre-industrial" time. To justify this method, individual models should be examined in their control-run modes to see whether they show significant drift (i.e., persistent, but unforced low-frequency changes) in climate patterns.

2.3.2.   Availability of suitable data

Regional climate models are driven primarily by lateral boundary conditions obtained from either reanalysis data or global climate models. In addition, if not produced prognostically, it is necessary to specify changes in the surface character (including man-made aspects like urbanization and irrigation, which may have significant regional-scale effects) and the fluxes from the surface to the atmosphere of radiatively or chemically active species, and changes in the free atmosphere of any radiatively active species that are used or modified internally by the RCM (such as greenhouse-gas concentrations, aerosols, and others). For relatively long-lived species, such as CO_2, CH_4, and N_2O, their abundances may be assumed to be constant through the troposphere. For some gases (e.g., CH_4) differences between the troposphere and stratosphere may have to be accounted for. For short-lived species (most importantly, aerosols, but possibly also tropospheric ozone) spatial variations should be considered.

Lateral boundary conditions must be sufficient to define both the instantaneous state of the atmosphere at the boundary and the fluxes across the boundary of the RCM's primary prognostic variables. Thus, one needs to specify boundary values for temperature, moisture, winds, geopotential height, and surface pressure. These coarse-resolution boundary values need to be interpolated horizontally and vertically from the driver grid to the grid of the RCM. For surface boundary conditions it may be necessary for the RCM to have a "spin up" period to stabilize the initial state (for example for soil moisture) relative to the model's atmosphere.

Boundary data must be specified on a 12-hourly or finer time scale (and interpolated to the time step of the RCM). A number of RCM experiments have been performed with 12-hourly boundary data (e.g., Qian et al. 2003) due to data availability constraints, but 6-hourly data are preferable. It is important for the RCM to be able to simulate diurnal variability. Poor simulations of the diurnal variability of precipitation are a serious problem, not only with almost all AOGCMs (Dai et al. 1999; Trenberth et al. 2003), but also with reanalysis data

(although the fidelity of the ECMWF ERA-40 reanalysis in this regard has not yet been tested, to my knowledge). In the United States, east of the Rockies, the onset of strong moist convection tends to occur earlier in the day in global models than it does in observations. An earlier onset of convection leads to models having generally weaker convection than is observed and, hence, smaller precipitation maxima. Problems like this might also occur in RCMs, so it is important, in RCM testing, to examine RCM simulations of the diurnal cycle of precipitation.

2.3.3.   Choosing a driver AOGCM

There are two factors that should be considered when choosing a driver AOGCM: (1) spanning the range of likely future changes in climate, and (2) the credibility of the driver model.

For temperature change, there is reasonable agreement between model projections of the patterns of change—the primary differences are associated with the overall magnitude of change, in turn related to the range of climate sensitivities exhibited by current models. Simply scaling results by global-mean temperature change can probably provide a reasonable estimate of the effects of sensitivity uncertainties.

For precipitation change, as already noted, there are large differences between models, with some models giving increases in precipitation over the California region and some giving decreases. The MAGICC/SCENGEN software package gives a way of quantifying uncertainties associated with these inter-model differences. The map below shows the inter-model signal-to-noise ratio (IM SNR) for Northern Hemisphere winter (DJF) precipitation, defined as the model-mean change divided by the inter-model standard deviation (using 17 models from the CMIP database). Note that the key criterion here is the absolute magnitude of the IM SNR, since its sign merely reflects the sign of the mean change.
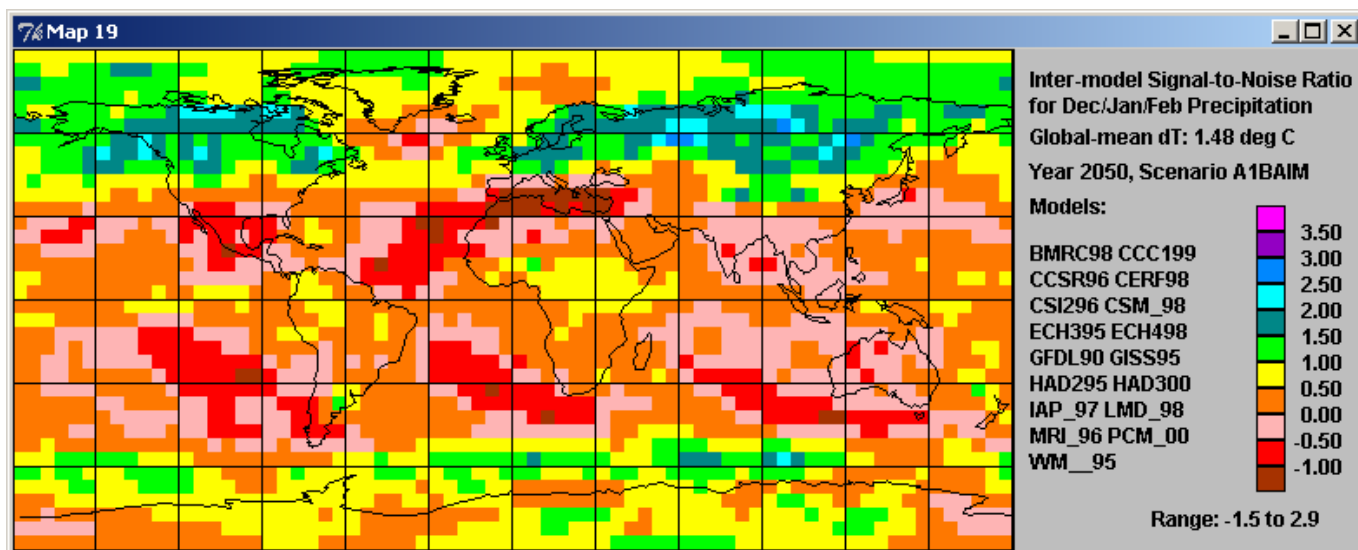

Figure 1: Inter-model signal-to-noise ratio for Northern Hemisphere winter (DJF) precipitation (based on 17 AOGCMs from the CMIP data set).

Ideally, one would like IM SNR values to be as large as possible, indicating that the model-mean signal (i.e., the projected magnitude of climate change) was much larger than the variability between models. From a statistical point of view, an IM SNR of at least 2 would be required to judge the mean change statistically significant. This is not the case for California. The zero change contour runs through southern California, so it is not surprising that IM SNR values are low over the whole California region. In fact, IM SNR values are less than 1 (indicating that the mean change is less than the inter-model uncertainty) for much of the globe. Over North America, IM SNR values above 1 occur only north of the Canada/U.S. border (and also for a small region around Montana). Since even the sign of future precipitation change is uncertain, it is important to choose a range of driver GCMs that spans both positive and negative changes.

The second factor that should be considered in choosing a driver GCM is the credibility of the model. The standard way to assess model credibility at the regional scale is by comparing model simulations of present-day climate with observed climate. Numerous papers have been published as products of the AMIP and CMIP model intercomparison projects that give the results of these comparisons (e.g., Covey et al. 2003). Choosing between models on the basis of these results is a difficult task, since models that perform well for one variable may not perform well for another (and vice versa). However, based on the set of models archived in the CMIP database, some models appear to be clearly superior. The superior models are the two Hadley Centre models (HadCM2 and HadCM3) and the Max Planck Institute model (ECHAM4) (see Appendix 1).

It should be noted, however, that there are more recent models that are under development (but not in the CMIP database) that need to be assessed. Most of this development work is directed towards producing results for use in the IPCC 4th Assessment Report. It should be further noted that these newer models need not be improvements over previous versions of the same model in terms of their ability to simulate present-day climate. For example, there are many aspects of HadCM2 that are superior to HadCM3. HadCM3, however, has improved physics and requires no flux adjustments in order to maintain a stable climate when run in control mode. A careful comparison of results over the RCM domain to be used for Californian climate simulations, including an assessment of the realism of ENSO simulations, would be required in order to choose between these two models as potential RCM drivers. (A related issue arising from the use of HadCM3 for RCM downscaling over Europe is discussed below.) For the new IPCC simulations, where model results will be available some time during 2004, an important consideration vis-à-vis their use as RCM drivers is resolution. Many of these new simulations will be at higher resolution than previous simulations (mainly T85, compared with T42, used earlier).

While the broad-brush assessment of AOGCMs that is given above provides a useful starting point, this assessment may not be entirely appropriate for deciding what models are most suitable for driving RCMs over the California study region.

First, it could be argued that a restricted area validation may be more suitable than considering the global performance of a model. Coquard et al. (2004) have performed such an evaluation for the California-Nevada region. They found that, among the CMIP models, HadCM3, HadCM2, and the CSIRO Mk2 model do relatively well at simulating near-surface temperature and precipitation.

Second, it is not clear which statistical metrics are best for AOGCM evaluation. Appendix 2 gives results for model bias (i.e., the difference between observed and model area-averages). Another common metric is the pattern correlation between observed and model climate. Models with lower bias are not necessarily the same as those with the highest pattern correlations. It might be argued that a pattern correlation metric is not suitable for AOGCM validation in the RCM-driver context, since it is at the pattern level that we expect the largest improvements to occur. However, higher pattern correlations over the RCM domain are likely to reflect the veracity of boundary values and fluxes (see below).

Third, as implied above, it is important to assess model performance, not only over the RCM domain, but also around the boundary of the domain. Whereas errors within the domain at the AOGCM scale may well be reduced at the RCM scale, this is more likely if the boundary values are of high quality. Such an assessment should evaluate AOGCM performance for both boundary values and boundary fluxes. I do not know of any published assessments of this type, but the Hadley Centre RCM (HadRM3) provides an interesting example. Over Europe, it is known that the global model HadCM3 has a large-scale bias in storm tracks. Since these tracks are important in defining fluxes across the western boundary of the RCM domain, any storm track error could seriously bias the RCM results. To reduce this problem, downscaling experiments with HadRM3 employ the intermediate-resolution atmospheric model HadAM3H as an intermediate step between the global AOGCM and the regional model.

Fourth, climate changes at the regional level occur both as a direct result of large-scale forcing and indirectly as a result of changes in natural modes of variability. In the California region, for example, ENSO variability has a strong influence on the large-scale patterns of precipitation, so it is important that any driver GCM is able to simulate ENSO variability realistically, and simulate the observed relationships between ENSO variability and regional climate. This does not mean that a driver model must be able to reproduce the history of past variations in ENSO—with a freely-varying AOGCM this is, of course, impossible. The key is to be able to simulate the statistical character of ENSO variability and the associated teleconnections. Note that, when driven by observed boundary condition data (i.e., reanalyses), RCMs are able to reproduce the small-scale patterns of ENSO-induced precipitation variability extremely well, and the detailed patterns of simulated precipitation in the RCM simulations are much superior to those for the driver GCM (see, e.g., Leung et al. 2003c). Such fidelity is dependent on driving the RCM with realistic boundary conditions, which is guaranteed when using reanalyses but which is far less certain when using an AOGCM as the driver—although even with less-than-perfect boundary conditions RCM simulations will almost certainly be superior to those for the driver GCM.

Whether driver models need to be able to simulate lower-frequency modes of variability, such as the Pacific Decadal Oscillation (PDO), is a moot point. As the historical record contains few manifestations of these low-frequency modes for which there is adequate observational data, we do not in general know enough about their statistical properties to be able to test model performance. Observed/model differences could reflect uncertain knowledge of the behavior of the real-world climate system. Furthermore, simulations of future changes in these low-frequency modes would have no relevance to their timing in the real world. As such, one could argue that simulated PDO changes (as an example of low-frequency variability) might be a disadvantage since, unless they could be confidently tied to the effects of external forcing, they would simply obfuscate any underlying signal.

2.3.4. Parameterization bias

A final source of potential error in RCM experiments is parameterization bias (Chen 2002). This problem arises because of differences between the driver GCM and the RCM in their physics parameterizations and numerical schemes, in turn arising because RCMs and GCMs are generally developed independently. The standard way to investigate this bias is to run the RCM driven by the GCM, but at the same resolution as the GCM. If this run is denoted RCM0, then the difference RCM0 minus GCM defines the bias (B = RCM0 - GCM). If B is applied to higher resolution RCM simulations as a correction factor, then these adjusted simulations tend to be more consistent with the driver GCM over the RCM domain (Chen 2002). Whether this improved consistency is a good or bad thing, however, remains uncertain.

2.4. Aerosols

The pioneering work on incorporating the effects of aerosols into RCMs is that of Qian and collaborators (Qian and Giorgi 1999; Qian et al. 2001; Giorgi et al. 2002, 2003; Qian et al. 2003. These references are collectively referred to hereafter as Q**). Just as for global AOGCMs (e.g., compare Dai et al. 2001a and 2001b), there are two approaches: (1) to include a fully interactive aerosol chemistry model within the RCM driven by appropriate emissions, or (2) to prescribe the time-evolving 3D aerosol fields based on the results from another model. Both methods require a radiation code in the RCM that can account for and include the radiative properties of the aerosols. The method of using short-wave-equivalent surface albedo changes that was employed in some early GCM experiments is considered unacceptable.

The first RCM that included aerosols prognostically (i.e., predicted internally from emissions) was a modified version of NCAR's regional model, RegCM2 (Giorgi et al. 1993b,c) that employed an interactive sulfur cycle (see Q**). This model has been used to study both the direct and indirect effects of sulfate aerosols. In these experiments, non-sulfate aerosols (carbonaceous or "black carbon" aerosols only—often referred to as *soot*) were accounted for by scaling against sulfate concentration, a less than satisfactory approach (as well-realized by the authors). To my knowledge, the only other RCM with an interactive sulfur cycle is the Hadley Centre's HadRM3 (Frei et al. 2002) (See also www.met-office.gov.uk/research/hadleycentre/ models/PRECIS), a development of HadRM2 (Jones et al. 1995, 1997; Noguer et al. 1998).

The other approach, prescribing the time-evolving 3D aerosol fields based on the results from another model, has been used by Qian et al. (2003) in experiments simulating present-day climate over East Asia. This approach has the advantage that many different types of aerosols can be included—in the Qian et al. experiments, sulfate, organic carbon, black carbon, mineral dust, sea salt and MSA (methane sulfonic acid) were considered. Aerosol fields were derived "off line" from the global aerosol/chemistry/transport model, MIRAGE, which couples together the PNNL global chemistry model GChM (Luecken et al. 1991; Benkovitz et al. 1994) and NCAR's CCM2.

The inclusion of black carbon (i.e., carbonaceous aerosols, or soot) is relatively new and potentially important, but the subject of considerable uncertainty. Estimates of the magnitude of black carbon forcing vary widely. Black carbon loadings are much higher over parts of Asia than in other parts of the world, and it is likely that black carbon influences over California are small. A quantitative assessment is required before any statement can be made about whether black carbon should be considered in RCM studies over California—especially given the lack of future scenarios for changes in black carbon loadings.

The experimental procedure used in MIRAGE is of interest, because it shows how complex experiments that employ separate aerosol and climate models can be. In MIRAGE, CCM2 provides cloud, precipitation, temperature, and circulation data to GChM (the latter two variables being continuously "nudged" back toward "observed" reanalysis data from the ECMWF ERA-15 reanalysis), and GChM employs these data along with emissions data to determine aerosol loadings, which are in turn passed to CCM2 to be used for radiative balance and cloud droplet calculations. Since MIRAGE tends to give unrealistic relative humidities, and because aerosol growth and optical properties are strongly influenced by relative humidity, aerosol growth was calculated in the RCM using the more realistic RCM humidities.

The message here is that there are important two-way interactions between aerosols and climate, and choices may have to be made about which aerosol characteristics will be calculated off-line and which will be calculated within the model. In practice, however, there may be no choices possible, since what can be done where depends on the type of off-line data available and on the capabilities of the RCM. In the Qian et al. (2001) study, the RCM has quite sophisticated aerosol physics to relate aerosol size distribution, total loading, and number density; to model the hygroscopic growth of aerosols; and to quantify aerosol optical properties. Completeness of processes and avoidance of duplication are issues that need to be addressed case by case.

A related issue is the diagnostic outputs of the RCM simulations. With aerosol simulations it is important to quantify the changing spatial patterns of aerosol radiative forcing (i.e., the net downward radiative flux at the top of the atmosphere), which can be compared with observations and/or across models. To do this "online" requires that the radiative calculations be made twice—once with aerosol scattering and absorption included, and once with aerosol effects ignored.

3.0   STATISTICAL DOWNSCALING

Although the basic methods of statistical downscaling have been used in meteorology for many years (see Wigley et al. 1990 for references), it was not until 1984 that the ideas were considered in a climate context (Kim et al. 1984). Modern regression-based methods were first applied by Wigley et al. (1990) and Karl et al. (1990), the former using Principal Components (PC) regression, and the latter using Canonical Correlation Analysis (CCA). Many different methods have been used since then, and the current state of the art has been reviewed in the IPCC Third Assessment Report (TAR) (Giorgi et al. 2001) and by Leung et al. (2003a). An extensive methods intercomparison project has been carried out by Wilby et al. (1998). This paper and Wilby and Wigley (1997) provide reviews of the most commonly used methods.

The basic goal of statistical downscaling is to derive a transfer function (i.e., a statistical relationship) between large-scale GCM data as predictors and small-scale climate variables as predictands. Either monthly mean or daily data may be used. Here, *large-scale* generally means from one to nine grid boxes, while *small-scale* often refers to a specific site. Downscaling transfer functions have been derived using a variety of statistical regression techniques. The use of PCs is common as a way of accounting for intercorrelations between predictor variables, or to capture spatial patterns in predictors and predictands in an efficient manner. In PC regression (as in Wigley et al. 1990) only the predictors are expressed as PCs. An extension of this is to express both predictors and predictands as PCs (e.g., Wigley and Tu Qipu 1983; Jones et al. 1987). A further extension is Canonical Correlation Analysis (CCA). PC-PC regression and CCA have been shown to produce very similar results, with the former method allowing a more comprehensive diagnosis of the relationships between variables (Cook et al. 1994).

A recent example of CCA is the work of Gershunov and Cayan (2003), of interest primarily because it examines the frequency of extreme precipitation events. This work does not give details of the strength of relationships at the individual station level, but their correlation maps (their Figure 1) show station-level correlations for downscaled frequencies of extreme daily precipitation that are generally below 0.5—statistically significant, but not high. The methods used warrant further investigation in the statistical downscaling context.

The basic assumptions of statistical downscaling are that the predictors are reliable, and that the derived transfer functions are stable in time (so that they may be applied to future situations). An initial set of predictor variables is usually chosen based on the following criteria: an assessment (often ad hoc) of which GCM variables are most reliable (which normally precludes GCM precipitation as a possible predictor); physical grounds and local climatological knowledge of what the controls are for a particular predictand; and preliminary statistical analyses. The number of predictors may subsequently be reduced based on statistical considerations (e.g., as in step-wise regression). The final predictor variables will depend on the predictand, the above considerations, and the study region.

There are two key differences between statistical and dynamical downscaling, related to their inputs and outputs. Consider the inputs first. In statistical downscaling, the input (predictor) variables are generally propinquitous with the predictands (i.e., the sites for the output

variables overlap with the input domain and the input domain for the downscaling is generally much smaller than for dynamical downscaling). Thus, one must assume that the driver GCM data are reliable on a relatively small scale, on the order of one to ten grid boxes. For dynamical downscaling, in the case where a study region constitutes only a small fraction of the model domain, the primary inputs (i.e., the lateral boundary conditions) are remote from the predictands and reflect much larger scale GCM information than is the case for statistical downscaling. Although these larger-scale data are more likely to be credible than the smaller-scale drivers for statistical downscaling, this merely transfers the driver reliability issue to the RCM itself, in the sense that it is the RCM outputs that are the drivers for any impacts model.

Key differences also exist for the predictands. In statistical downscaling the predictands are usually specific to a single site. For many impacts applications, separate analyses must be carried out for a range of sites within the study region. This may lead to site-to-site spatial relationships that are unrealistic. Canonical correlation methods (e.g., Karl et al. 1990) or the use of principal components as predictands may overcome this problem, but these methods are rarely applied. Dynamical downscaling avoids this problem by producing the complete spatial field over and beyond any specific study region.

In addition to spatial correlations, inter-variable correlations may be a problem. Statistical downscaling methods generally produce relationships for each climate variable independently, and so may lead to incorrect inter-variable relationships. In principle, this factor should not be a problem with dynamical downscaling—provided the physics of the RCM is perfect. Of course, model physics is not perfect, so for both downscaling methods the credibility of inter-variable relationships should be tested.

With statistical downscaling, the inputs used (i.e., the predictors) depend on the specific predictand. Thus, the regression relationships for different predictands may be quite dissimilar. Nevertheless, a fairly standard set of potential predictors has evolved with experience. The predictor variables are generally specified for the GCM grid box in which the analysis site is located (although, for predictors involving spatial gradients, calculation of the grid-box value may involve using data from surrounding grid boxes).

At the surface, the most common predictors are: mean sea level pressure (MSLP); north-south and east-west MSLP gradients (or wind speeds); surface divergence; and surface vorticity. Above the surface, the predictors are: a moisture variable representing the surface to 700hPa layer; 500hPa geopotential height and north-south and east-west gradients (or wind speeds); 500hPa divergence; and 500hPa vorticity. When analyses are carried out on a daily timescale, it is common practice for precipitation to use lagged as well as contemporary predictors. Note that GCM-based, grid-box temperature and precipitation are not used, since, at the grid-box level, these are considered less reliable than the above-listed predictors.

In addition, remote predictors such as an ENSO index or sea-surface temperatures (see, e.g., Katz and Parlange 1993) may be used in specific cases. As a final point, the assumption that the best predictors are those for the grid box that contains the analysis site is not necessarily correct. Wilby and Wigley (2000) found that maximum site/predictor correlations often occurred with

slightly displaced predictors (as might be expected because of transport in, for example, westerly flow regimes), so this possibility should be explored.

In any regression-based procedure, an important aspect is model validation. Ideally, regression equations derived over a calibration interval should be tested over an independent validation interval. It is true that confidence intervals for overall explained variance and individual regression coefficients can be quantified directly as part of the calibration, but the reliability of these estimates is questionable in statistical downscaling exercises, primarily because the predictor variables may be significantly correlated (leading to the problem of multi-collinearity), and because the predictors and predictands often show strong temporal autocorrelation. These problems may invalidate the assumptions underlying theoretical confidence limit calculations, and may lead to spurious results. If insufficient data are available for independent validation, then an alternative is to use sub-sample replication.

In regression-based statistical downscaling, the derived regression equations only explain a fraction of the predictand variance; generally more for temperature than for precipitation as a predictand. There are two reasons for this. The first is because a fraction of the predictand variance (N%, say) is likely to be the result of inherently unpredictable "weather noise". The second is that the regression equation may, for statistical or structural reasons, simply be incapable of explaining all of the potentially predictable component (the (100–N)%). This is both an advantage and a disadvantage. If the regression equation were perfect (i.e., able to explain all of the potentially-predictable variance), then stochastic simulation of the residual weather noise may be used to simulate multiple realizations of the future and so produce a quantitative assessment of uncertainties. Of course, to do this realistically one must use a noise component that has the correct statistical time series and spatial correlation structure. This is a non-trivial problem (see, e.g., Wilks 1998; Wilks and Wilby 1999). Temporal and spatial noise realism has rarely been considered in any practical applications of statistical downscaling (but see Wilby et al. 2003).

A more important problem, at least for precipitation, is that the derived regression equation may fail to explain all of the potentially predictable variance. Exploration of a range of predictor variables and independent validation of calibrated regressions can provide some insights into the magnitude of this problem, but it remains a poorly understood issue. A related problem here is the stability of the regression equation. Is an equation derived using, say, late twentieth-century data applicable to the late twenty-first century, or will the predictor/predictand relationships change with time? If the relationships have a sound physical basis, then they are clearly more likely to remain stable in time; indeed, this has been demonstrated by Charles et al. (1999) and Wilby and Wigley (2000).

4.0   COMPARISONS OF DOWNSCALING METHODS

There have been a number of studies comparing dynamical downscaling using RCMs and statistical downscaling (e.g., Kidson and Thompson 1998; Murphy 1999, 2000; Hellström et al. 2001; and Hanssen-Bauer 2003. See also Giorgi et al. 2001 and Yarnal et al. 2001). Specific impacts-related papers include Wilby et al. (1999, 2000), Wilby and Dettinger (2000), and Hay and Clark (2003). These latter four papers deal with runoff estimation for the San Juan River (CO), the Animas River (CO), three rivers with headwaters in the Sierra Nevada, and the Animas, Carson (NV), and Cle Elum (WA) rivers, respectively. The general conclusion of these comparisons is that both methods give comparable results in terms of their skill in reproducing the mean and variability of present-day climatic or river flow conditions, on both daily and monthly time scales. One should note, however, that runoff simulations in snowpack-dominated basins smooth out errors in the daily sequencing of precipitation events through moisture storage in the snowpack.

It should be noted, however, that, in hydrological applications where snowmelt is a significant contributor to runoff, differences between RCM and real-world orography, even at the relatively fine resolution of current RCMs (around 50km), still require empirical corrections to be applied to RCM results. Furthermore, in spite of comparable performance when simulating the present, the two methods can give quite different results when applied to future conditions, for reasons that are not yet clear. Hay and Clark (2003) state "Future work is warranted to identify the causes for (and removal of) systematic biases in [dynamically downscaled] simulations of daily variability in local climate. Until then, [statistically-based] simulations of runoff appear to be the safer downscaling choice."

Global downscaling using high-resolution atmospheric models is a new approach, and thus has received relatively little formal evaluation. In this approach, a high-resolution global atmospheric model is forced by sea-surface temperatures and sea-ice extents from a coarse-resolution AOGCM. As discussed in Appendix 3, this downscaling approach avoids numerical and other problems associated with nesting one model within another. As a result, regional-scale results of high-resolution global models can rival those of nested RCMs (see Appendix 3 for examples). The computational expense of the global models is much greater, however. Duffy et al. (2003) evaluated simulations of this type performed with the CCM3 model and showed that the large-scale ($\geq$ 300 km) climate features at T170 (~75 km resolution) and T239 (~50 km resolution) agree much more closely with a broad range of observations than do results from the same model at T42 (~300 km). This is true even when the high-resolution results are filtered to admit scales only at the T42 grid size and larger. This means that the high-resolution simulations are not only more detailed, but also more accurate on large scales, than a coarse-resolution simulation with the same model. This suggests that these high-resolution simulations should provide superior lateral boundary conditions for driving nested RCMs.

Although there is no clear difference in performance between statistical downscaling and RCM-based dynamical downscaling, there are other factors that represent relative advantages and disadvantages of the two methods.

1. Both methods rely on the credibility of the driver GCM data, but at different spatial scales. For statistical downscaling, the dependence is more direct, because it is local grid-box data that are generally used as drivers, and it is GCM reliability at the 1 to 10 grid-box scale that is important. For dynamical downscaling, it is larger-scale aspects of the driver GCM that are the controls. Furthermore, GCM errors at the lateral boundaries may be mitigated somewhat as they are propagated from the boundaries to the interior of the domain, especially if the higher resolution of the RCM is echoed by improved physics. For global downscaling using a high-resolution AGCM, only the SSTs and sea ice distribution from the coarse-resolution AOGCM are relied upon.
2. It is often assumed a priori that physically-based methods are superior to more empirical methods. At the present state of their development, however, there are still important approximations involved in the physics parameterizations of RCMs. The scope for improvement of RCM-based downscaling is therefore greater than for statistical downscaling, but the time scale over which this improvement might occur may be many years.
3. A disadvantage of RCM-based downscaling is that it is highly computer intensive. This problem is exacerbated because multiple realizations may be required to define the signal—not only multiple realizations of the RCM for given boundary conditions, but also multiple realizations of the boundary conditions based on different GCM ensemble members. RCM studies require days of wall-clock time on today's fastest computers for each realization. Statistical downscaling may also suffer from noise at the GCM driver level. However, statistical downscaling experiments can be carried out in seconds on a personal computer (producing multiple realizations of the weather noise if required), giving much more scope for the assessment of uncertainties.

The bottom line here is that statistical and dynamical downscaling (whether with a nested RCM or a high-resolution global model) complement each other, and, where possible, both should be carried out in parallel.

5.0 REFERENCES

Benkovitz, C. M., Berkowitz, C. M., Easter, R. C., Nemesure, S., Wagener, R. and Schwartz, S. E., 1994: Sulfate over the North Atlantic and adjacent continental areas: Evaluation for October and November, 1996, using a three-dimensional model driven by observation-derived meteorology. *Journal of Geophysical Research* 99, 20725–20756.

Cayan, D. R., 1996: Interannual climate variability and snowpack in the western United States. *Journal of Climate* 9, 928–948.

Cayan, D. R. and Roads, J. O., 1984: Local relationships between United States west coast precipitation and monthly mean circulation parameters. *Monthly Weather Review* 112, 1276–1282.

Cayan, D. R., Redmond, K. T. and Riddle, L.G., 1999: ENSO and hydrologic extremes in the western United States. *Journal of Climate* 12, 2881–2893.

Charles, S. P., Bates, B. C., Whetton, P. H. and Hughes, J. P., 1999: Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Climate Research* 12, 1–14.

Chen, S. -C., 2002: Model mismatch between global and regional simulations. *Geophysical Research Letters* 29, 10.1029/2001GL013570, 4-1–4-4.

Chen, S. -C., Roads, J. O., Juang, H. -M. and Kanamitsu, M., 1999: Global to regional modeling: California wintertime precipitation. *Journal of Geophysical Research* 104, 31517–31532.

Cook, E. R., Briffa, K. R. and Jones, P. D., 1994: Spatial regression methods in dendroclimatology: A review and comparison of two techniques. *International Journal of Climatology* 14, 379–402.

Coquard, J., P. B. Duffy, and K. E. Taylor, 2004: Present and future surface climate in the western U.S. as simulated by 15 global climate models. Submitted to *Climate Dynamics.*

Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P. D., Lambert, S. J., Mann, M. E., Phillips, T. J. and Taylor, K. E., 2003: An overview of results from the Coupled Model Intercomparison Project (CMIP). *Global and Planetary Change* 37, 103–133.

Dai, A., Giorgi, F. and Trenberth, K. E., 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *Journal of Geophysical Research* 104, 6377–6402.

Dai, A., Wigley, T. M. L., Boville, B., Kiehl, J. T. and Buja, L., 2001a: Climates of the twentieth and twenty-first centuries simulated by the NCAR Climate System Model. *Journal of Climate* 14, 485–519.

Dai, A., Meehl, G. A., Washington, W. M., Wigley, T. M. L. and Arblaster, J. M., 2001b: Ensemble simulation of twenty-first century climate changes: Business as usual vs. CO_2 stabilization. *Bulletin of the American Meteorological Society* 82, 2377–2388.

Denis, B., Laprise, R., Cote, J. and Cayan, D., 2002: Downscaling ability of one-way nested regional climate models: The big-brother experiment. *Climate Dynamics* 18, 627–646.

Denis, B., Laprise, R. and Cayan, D., 2003: Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. *Climate Dynamics* 20, 107–126.

Dickinson, R. E., Errico, R. M., Giorgi, F. and Bates, G. T., 1989: A regional climate model for the western United States. *Climatic Change* 15, 383–422.

Duffy, P. B., B. Govindasamy, J. Iorio, J. Milovich, K. Sperber, K. E. Taylor, M. F. Wehner, and S. L. Thompson, 2003: High resolution simulations of global climate, part 1: present climate, *Climate Dynamics* 21, 371–390.

Frei, C., Christensen, J. H., Déqué, M., Jacob, D., Jones, R. G. and Vidale, P. L., 2002: Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. *Journal of Geophysical Research* 108(D3), 4124, doi:10.1029/2002JD002287.

Gershunov, A. and Cayan, D. R., 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *Journal of Climate* 16, 2753–2765.

Gibson, J. K., et al., 1997: ERA Description. *ECMWF Re-Analysis Project Report Series No. 1.* ECMWF, Reading, U.K. 66pp.

Giorgi, F. and Bates, G. T., 1989: On the climatological skill of a regional model over complex terrain. *Monthly Weather Review* 117, 2325–2347.

Giorgi, F., Bates, G. T. and Nieman, S. J., 1993a: The multiyear surface climatology of a regional atmospheric model over the western United States. *Journal of Climate* 6, 75–95.

Giorgi, F., Marinucci, M. R. and Bates, G. T., 1993b: Development of a second generation regional climate model (RegCM2). Part I: Boundary-layer and radiative transfer processes. *Monthly Weather Review* 121, 2794–2813.

Giorgi, F., Marinucci, M. R., Bates, G. T. and De Canio, G., 1993c: Development of a second generation regional climate model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions. *Monthly Weather Review* 121, 2814–2832.

Giorgi, F., Bi, X. Q. and Qian, Y., 2002: Direct radiative forcing and regional climate effects of anthropogenic aerosols over East Asia: a regional coupled climate-chemistry/aerosol model study. *Journal of Geophysical Research* 107(D20), 4439, doi:10.1029/2001JD001066.

Giorgi, F., Bi, X. Q. and Qian, Y., 2003: Indirect vs. direct effects of anthropogenic sulfate on the climate of East Asia as simulated with a regional coupled climate-chemistry/aerosol model. *Climatic Change* 58, 345–376.

Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., Von Storch, H., Whetton, P., Jones, R., Mearns, L. and Fu, C., 2001: Regional climate information – Evaluation and projections. (In) *Climate Change 2001: The Scientific Basis* (Eds. J.T. Houghton, et al.). Cambridge University Press, Cambridge, pp. 583–638.

Govindasamy, B., Duffy, P. B. and Coquard, J. 2003: High resolution simulations of global climate, part 2: effects of increased greenhouse gases. *Climate Dynamics* 21, 391–404.

Hanssen-Bauer, I., Førland, E. J., Haugen, J. E. and Tveito, O. E., 2003: Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. *Climate Research* 25, 15–27.

Hay, L. E. and Clark, M. P., 2003: Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *Journal of Hydrology* 282, 56–75.

Hellström, C., Chen., Achberger, C. and Räisänen, J., 2001: Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. *Climate Research* 19, 45–55.

Iorio, J. P., Duffy, P. B., Govindasamy, B. and Thompson, S. L., 2004: Effects of model resolution and subgrid-scale physics on the simulation of daily precipitation in the continental United States. Submitted to *Climate Dynamics.*

Jones, P. D., Wigley, T. M. L. and Briffa, K. R., 1987: *Monthly Mean Pressure Reconstructions for Europe (Back to 1780) and North America (to 1858).* U.S. Department of Energy, Carbon Dioxide Research Division, Technical Report TR037, 99 pp.

Jones, R. G., Murphy, J. M. and Noguer, M., 1995: Simulation of climate change over Europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries, *Quarterly Journal of the Royal Meteorological Society* 121, 1413–1449.

Jones, R.G., Murphy, J.M., Noguer, M. and Keen, A.B., 1997: Simulation of climate change over Europe using a nested regional-climate model. II: Comparison of driving and regional model responses to a doubling of carbon dioxide, *Quarterly Journal of the Royal Meteorological Society* 123, 265–292.

Kalnay, E., et al., 1996: The NCAR/NCEP 40-year reanalysis project. *Bulletin of the American Meteorological Society* 77, 437–471.

Karl, T. R., Wang, W. C., Schlesinger, M. E., Knight, R. W. and Portman, D., 1990: A method of relating general circulation model simulated climate to the observed climate, Part I: Seasonal statistics. *Journal of Climate* 3, 1053–1079.

Katz, R. W. and Parlange, M. B., 1993: Effects of an index of atmospheric circulation on stochastic properties of precipitation. *Water Resources Research* 29, 2335–2344.

Kidson, J. W. and Thompson, C. S., 1998: Comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate* 11, 735–753.

Kim, J., 1997: Precipitation and snow budget over the southwestern United States during the 1994–1995 winter season in a mesoscale simulation. *Water Resources Research* 33, 2831–2839.

Kim, J., Miller, N. L., Farrara, J. D. and Hong, S. -Y., 2000: A seasonal precipitation and streamflow hindcast and prediction study in the western United States during the 1997/98 winter season using a dynamic downscaling system. *Journal of Hydrometeorology* 3, 322–334.

Kim, J. W., Chang, J. T., Baker, N. L., Wilks, D. S. and Gates, W. L., 1984: The statistical problem of climate inversion: Determination of the relationship between local and large-scale climate. *Monthly Weather Review* 112, 2069–2077.

Leung, R. L. and Ghan, S. J., 1999a: Pacific Northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part I: control simulations. *Journal of Climate* 12, 2010–2030.

Leung, R. L. and Ghan, S. J., 1999b: Pacific Northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part II: 2xCO$_2$ simulations. *Journal of Climate* 12, 2031–2053.

Leung, R. L., Mearns, L. O., Giorgi, F. and Wilby, R. L., 2003a: Regional climate research: Needs and opportunities. *Bulletin of the American Meteorological Society* 84, 89–95.

Leung, R. L., Qian, Y. and Bian, X., 2003b: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part I: Seasonal statistics. *Journal of Climate* 16, 1892–1911.

Leung, R. L., Qian, Y., Bian, X. and Hunt, A., 2003c: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part II: Mesoscale ENSO anomalies. *Journal of Climate* 16, 1912–1928.

Luecken, D. J., Berkowitz, C. M. and Easter, R. C., 1991: Use of a three-dimensional cloud chemistry model to study the transatlantic transport of soluble sulfur species. *Journal of Geophysical Research* 96, 22477–22490.

Mantua, N. J. and Hare, S. R., 2002: The Pacific Decadal Oscillation. *Journal of Oceanography* 58, 35–44.

Mitchell, T. P. and Blier, W., 1997: The variability of wintertime precipitation in the region of California. *Journal of Climate* 10, 2261–2276.

Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate* 12, 2256–2284.

Murphy, J., 2000: Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *International Journal of Climatology* 20, 489–501.

Noguer, M., Jones, R. G. and Murphy, J. M., 1998: Sources of systematic errors in the climatology of a nested Regional Climate Model (RCM) over Europe. *Climate Dynamics* 14, 691–712.

Qian, Y. and Giorgi, F., 1999: Interactive coupling of regional climate and sulfate aerosol models over eastern Asia. *Journal of Geophysical Research* 104, 6477–6499.

Qian, Y., Giorgi, F., Huang, Y., Chameides, W. L. and Luo, C., 2001: Regional simulation of anthropogenic sulfur over East Asia and its sensitivity to model parameters. *Tellus* 53B, 171–191.

Qian, Y., Leung, R. L., Ghan, S. J. and Giorgi, F., 2003: Regional effects of aerosols over China: modeling and observation. *Tellus* 55B, 914–934.

Santer, B. D. et al., 2004: Influence of satellite data uncertainties on the detection of externally-forced climate change. Submitted for publication.

Santer, B. D., Wigley, T. M. L., Schlesinger, M. E. and Mitchell, J. F. B., 1990: *Developing Climate Scenarios from Equilibrium GCM Results.* Max-Planck-Institut für Meteorologie Report No. 47, Hamburg, Germany, 29 pp.

Simmons, A. J. and Gibson, J. K., 2003: The ERA-40 Project Plan. *ERA-40 Project Report Series No.1.* ECMWF, Reading, U.K., 62pp.

Takle, E.S., Gutowski, W. J. and 18 others, 1999: Project to intercompare regional climate simulations (PIRCS): description and initial results. *Journal of Geophysical Research* 104, 19443–19461.

Trenberth, K. E., Dai, A., Rasmussen, R. M. and Parsons, D. B., 2003: The changing character of precipitation. *Bulletin of the American Meteorological Society* 84, 1205–1217.

Trenberth, K. E. and Guillemot, C. J., 1996: Physical processes involved in the 1998 drought and 1993 floods in North America. *Journal of Climate* 9, 1288–1298.
.
Trenberth, K. E. and Hurrell, J. W., 1994: Decadal atmosphere-ocean variations in the Pacific. *Climate Dynamics* 9, 303–319.

Wigley, T. M. L. and Tu Qipu, 1983: Crop-climate modelling using spatial patterns of yield and climate: Part 1, Background and an example from Australia. *Journal of Climate and Applied Meteorology* 22, 1831–1841.

Wigley, T. M. L., Jones, P. D., Briffa, K. R. and Smith, G., 1990: Obtaining sub-grid scale information from coarse-resolution general circulation model output. *Journal of Geophysical Research* 95, 1943–1953.

Wilby, R. L. and Dettinger, M. D., 2000: Streamflow changes in the Sierra Nevada, California, simulated using a statistically downscaled General Circulation Model scenario of climate change. (In) *Linking Climate Change to Land Surface Change* (Eds. S. J. McLaren and D. R. Kniveton). Kluwer Academic Publishers, Netherlands, pp. 99-121.

Wilby, R. L. and Wigley, T. M. L., 1997: Downscaling general circulation model output: A review of methods and limitations. *Progress in Physical Geography* 21, 530–548.

Wilby, R. L. and Wigley, T. M. L., 2000: Precipitation predictors for downscaling: observed and General Circulation Model relationships. *International Journal of Climatology* 20, 641–661.

Wilby, R. L. and Wigley, T. M. L., 2002: Future changes in the distribution of daily precipitation totals across North America. *Geophysical Research Letters* 29, 10.1029/2001GL013048.

Wilby, R. L., Hay, L. E., Gutowski, W. J., Arritt, R. W., Takle, E. S., Pan, Z., Leavesley, G. H. and Clark, M. P., 2000: Hydrological response to dynamically and statistically downscaled climate model output. *Geophysical Research Letters* 27, 1199–1202.

Wilby, R. L., Hay, L. E. and Leavesley, G. H., 1999: A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *Journal of Hydrology* 225, 67–91.

Wilby, R. L., Tomlinson, O. J. and Dawson, C. W., 2003: Multi-site simulation of precipitation by conditional resampling. *Climate Research* 23, 183–194.

Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J. and Wilks, D. S., 1998: Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research* 34, 2995–3008.

Wilks, D. S., 1998: Multisite generalization of a daily stochastic precipitation model. *Journal of Hydrology* 210, 178–191.

Wilks, D. S. and Wilby, R. L., 1999: The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* 23, 329–357.

Xie, P. and Arkin, P. A., 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society* 78, 2539–2558.

Yarnal, B., Comrie, A. C., Frakes, B. and Brown, D. P., 2001: Review. Developments and prospects in synoptic climatology. *International Journal of Climatology* 21, 1923–1950.

Appendix 1
Signal-to-Noise Ratios

An important issue is whether or not there is likely to be a statistically significant climate-change signal over California due to future anthropogenic forcing. To provide insight into this, I have used results from the MAGICC/SCENGEN software package (downloadable from www.cgd.ucar.edu). The results below are for Northern Hemisphere winter (Dec., Jan., Feb.) precipitation under the A1B emissions scenario for a 20-year period centered about 2050. The central estimate for global-mean warming over 1990–2050 under this scenario is 1.48°C. AOGCM data are from the CMIP database. I show results for two Hadley Centre models, HadCM2 and HadCM3 (since these are arguably two of the best models available, and since HadCM2 was one of the two AOGCMs used in the U.S. National Assessment), and for the average over 17 models in the CMIP database. Signal-to-noise ratio (SNR) is defined as change divided by the control-run inter-annual standard deviation (with both statistics taken from the model results using 20-year sample sizes).

For HadCM2 (top panel), all of California shows an increase in DJF precipitation. Over most of the state, the SNR lies between 0.5 and 1.0. Thus, the climate change signal for this variable is actually smaller than the inter-annual variability. Moving to a later future date would increase the signal-to-noise ratio, so this would indicate that a time interval towards the end of the century would be required to produce meaningful results with HadCM2 as the driver model.

For HadCM3 (middle panel), almost the opposite results are obtained. All of California shows a decrease in DJF precipitation with SNR in the range of -0.5 to -1.0. Again, if HadCM3 is used as a driver model, then the driver data should come from a period towards the end of the century.

When SNR values are averaged over the full set of 17 models in the SCENGEN database (bottom panel), California shows a small increase in DJF precipitation everywhere except for the southernmost part of the state. The average SNR is less than 0.5 everywhere, indicating that there are no models that have large SNR values. In fact, if all 17 models are examined individually, the two Hadley Centre models are close to the lower and upper bounds for SNR.
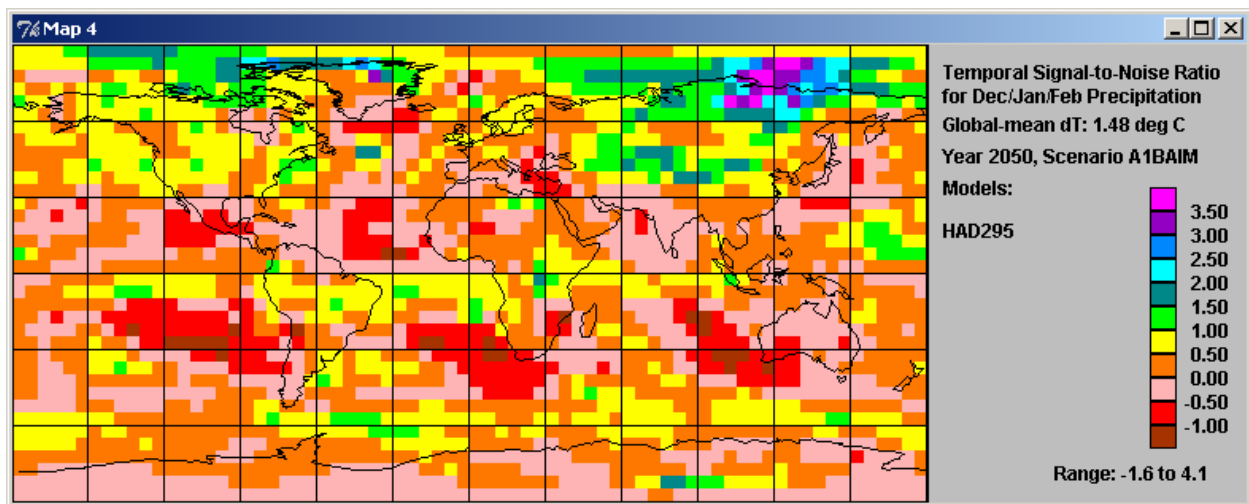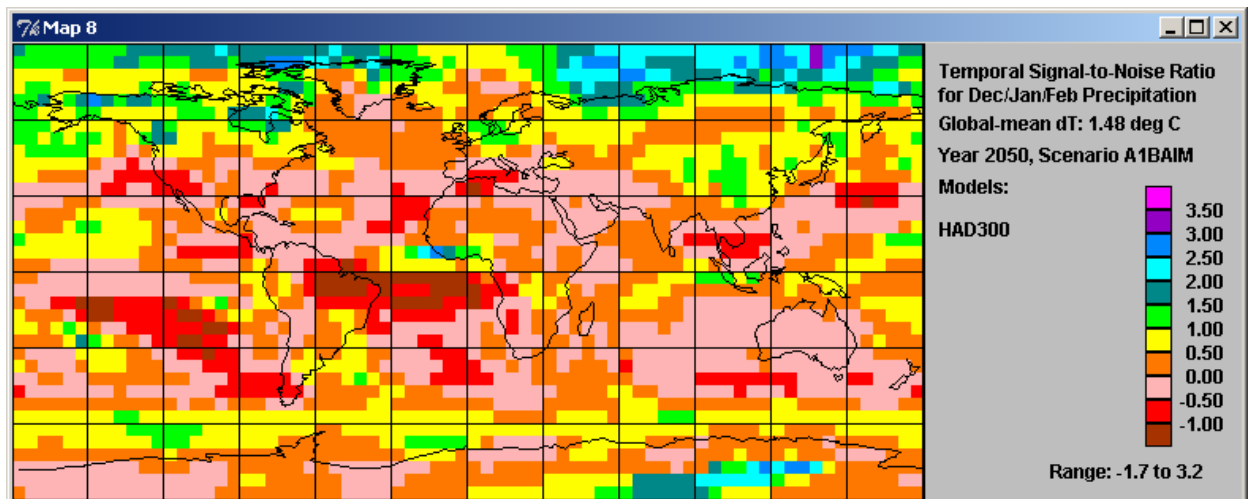
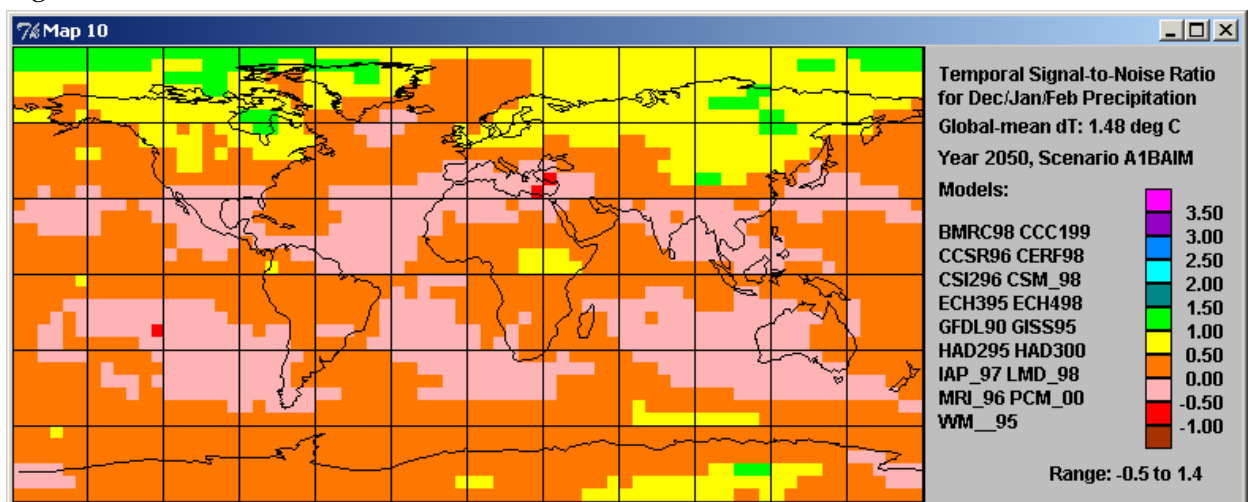Figure A-1. HadCM2 results.

Figure A-2. HadCM3 results.

Figure A-3. SCENGEN results.

Appendix 2

AOGCM Validation

Global climate models have deficiencies that are common to most models, particularly in simulating precipitation. Some of these are clearly related to the coarse resolution of the global models. The most obvious effect of resolution is on processes related to orography—in the western United States, a coarse resolution model replaces the mountains by much smoother and less extreme terrain variations. A second influence is through the representation of moisture sources and the circulation patterns that transport moisture from these sources to precipitation regions.

The map below shows the average AOGCM bias in Northern Hemisphere winter (DJF) precipitation. Model precipitation here is the average of control-run results from 17 models in the CMIP database, and the bias is model-minus-observed-precipitation expressed as a percentage relative to observed (1980–1999 mean) precipitation. Although the map shows an average result, the large overestimate of precipitation that is seen over the western United States is common to all models. It arises partly because the control-run forcings in the CMIP runs are not the same as present-day forcing, and partly because of the resolution-dependent issues noted above. There are also uncertainties in the gridded precipitation database that has been used for the comparison. Here I have used the CMAP (CPC Merged Analysis of Precipitation) data set. This data set uses the Xie/Arkin data, in which gaps are infilled using NCEP/NCAR reanalysis information (Xie and Arkin 1997). To my knowledge, there has been no comprehensive assessment of this apparent model bias.
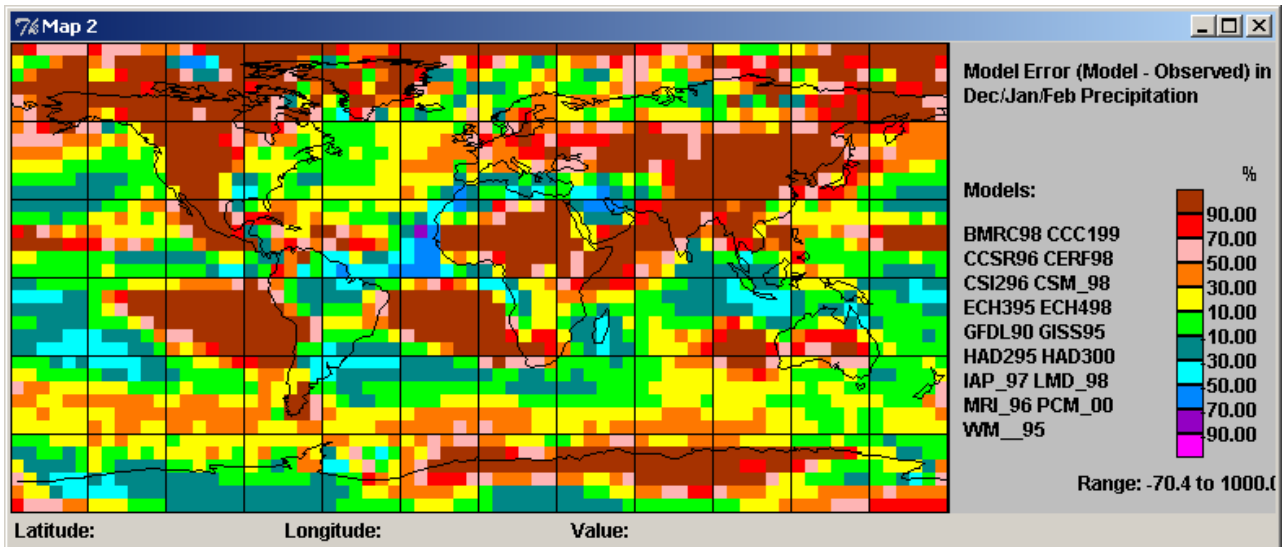

 Figure A-4. Average AOGCM bias in Northern Hemisphere winter (DJF) precipitation.

Regional climate models should improve on this representation of present-day precipitation, simply because they are able to represent the orography better. It should be noted, however, that baseline simulation errors need not lead to similar errors in projected changes. This is clear

from the fact that models with similar control-run errors can lead to quite different projected changes. In general, when comparing models, there appears to be no relationship between baseline error and projected precipitation change over the California region.

Appendix 3

Dynamical Downscaling Using High-resolution Global Atmospheric Models

The most common approach to simulating regional climate is to dynamically downscale the climate of a coarse-resolution AOGCM using a fine-resolution nested regional climate model (RCM). As discussed above, this approach has proven benefits, as well as important drawbacks. One drawback is that numerical problems associated with nesting one model within another can degrade the RCM solution. Another, more important issue is that the quality of the RCM solution will be degraded by deficiencies in the quality of the lateral boundary condition data that drive the simulation. Since these data are typically obtained from a coarse-resolution AOGCM, they often have important biases.

An alternative approach that avoids these problems is to perform dynamical downscaling using a high-resolution <u>global</u> atmospheric model. In this approach, the atmospheric model is driven by lower boundary condition data (sea-surface temperatures and sea-ice extents) from a coarse-resolution AOGCM. Because the climate is largely determined by these boundary condition data, the high-resolution atmospheric model is not predicting the climate, but rather downscaling the climate of the coarse-resolution AOGCM to fine resolution. This approach has been demonstrated most recently by Duffy et al. (2003) who performed high-resolution global simulations of the present climate, and by Govindasamy et al. (2003), who downscaled an increased greenhouse gas climate from the PCM model. In this downscaling approach there is no nesting of one model within another; thus the problems mentioned above associated with nested RCMs are avoided. As a result, this approach can produce regional solutions that are not only greatly superior to those from coarse-resolution global models, but also rival those of RCMs driven by coarse-resolution AOGCMs. The figures below show example results from recent high-resolution global AGCM simulations.

It should also be mentioned that the two dynamical downscaling approaches can be combined to produce ultra-high-resolution regional solutions. First, the coarse-resolution AOGCM solution is downscaled using a high-resolution global atmospheric mode; further downscaling is then performed using an ultra high-resolution nested RCM. In this hybrid approach, the problems associated with nesting models are minimized by driving the RCM with a high-resolution global model. This minimizes numerical problems at the edge of the RCM domain (because the resolution jump between the RCM and driving GCM is minimized). Also, the quality of the high-resolution lateral boundary condition data should in general be superior to that obtained from coarse-resolution models.
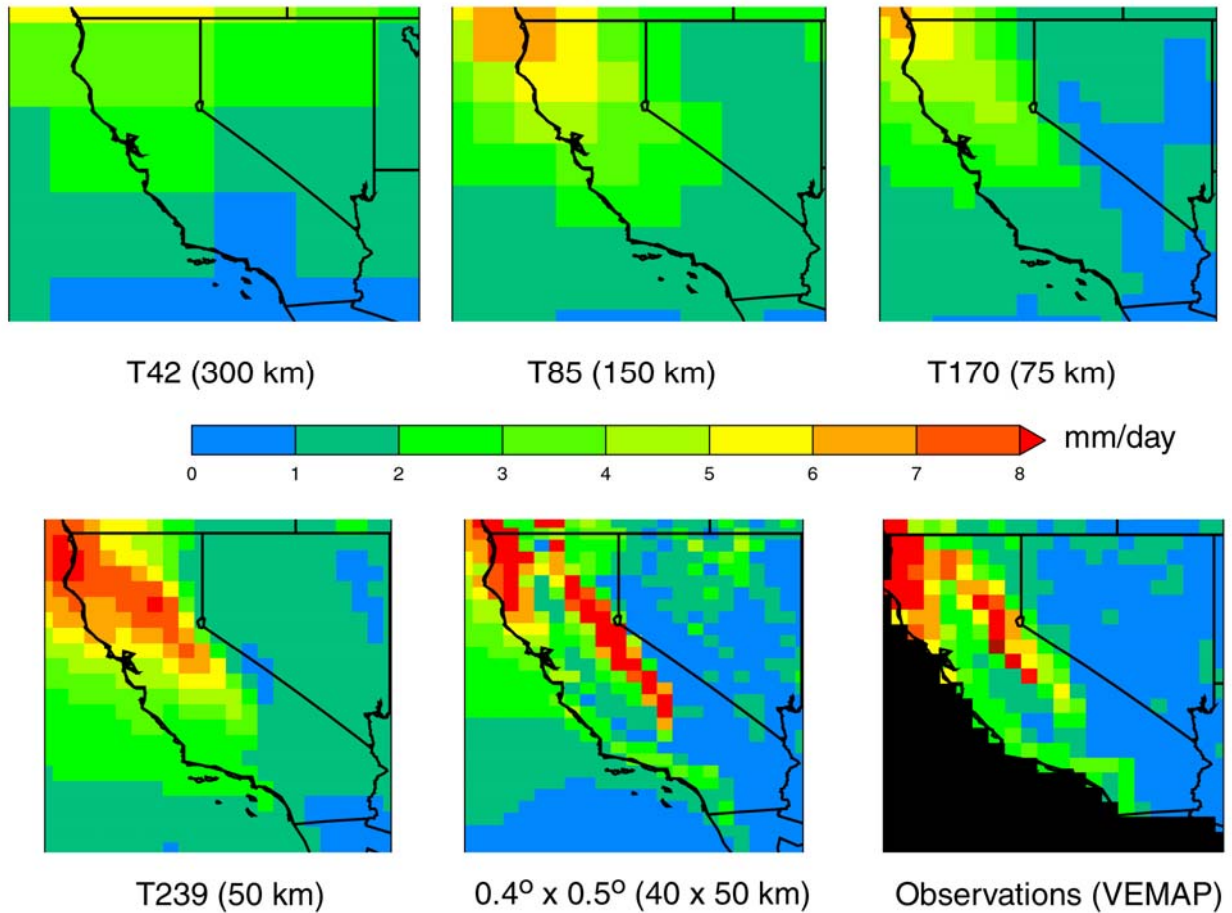
Figure A-5. Winter (DJF) precipitation in California and Nevada. Lower right panel is observed precipitation from VEMAP. The three upper panels and the lower left panel show results of the CCM3 global atmospheric model at increasingly fine spatial resolutions. The bottom middle panel shows results of the CAM2 global model at about 40 x 50 km resolution. This simulation uses the finite-volume dynamical core, which allows better representation of orography than is possible with the Eulerian spectral dynamics used in the other simulations. As a result, in this simulation the spatial pattern of precipitation (features such as the relative dryness of Nevada and of California's Central Valley) are more accurately represented.
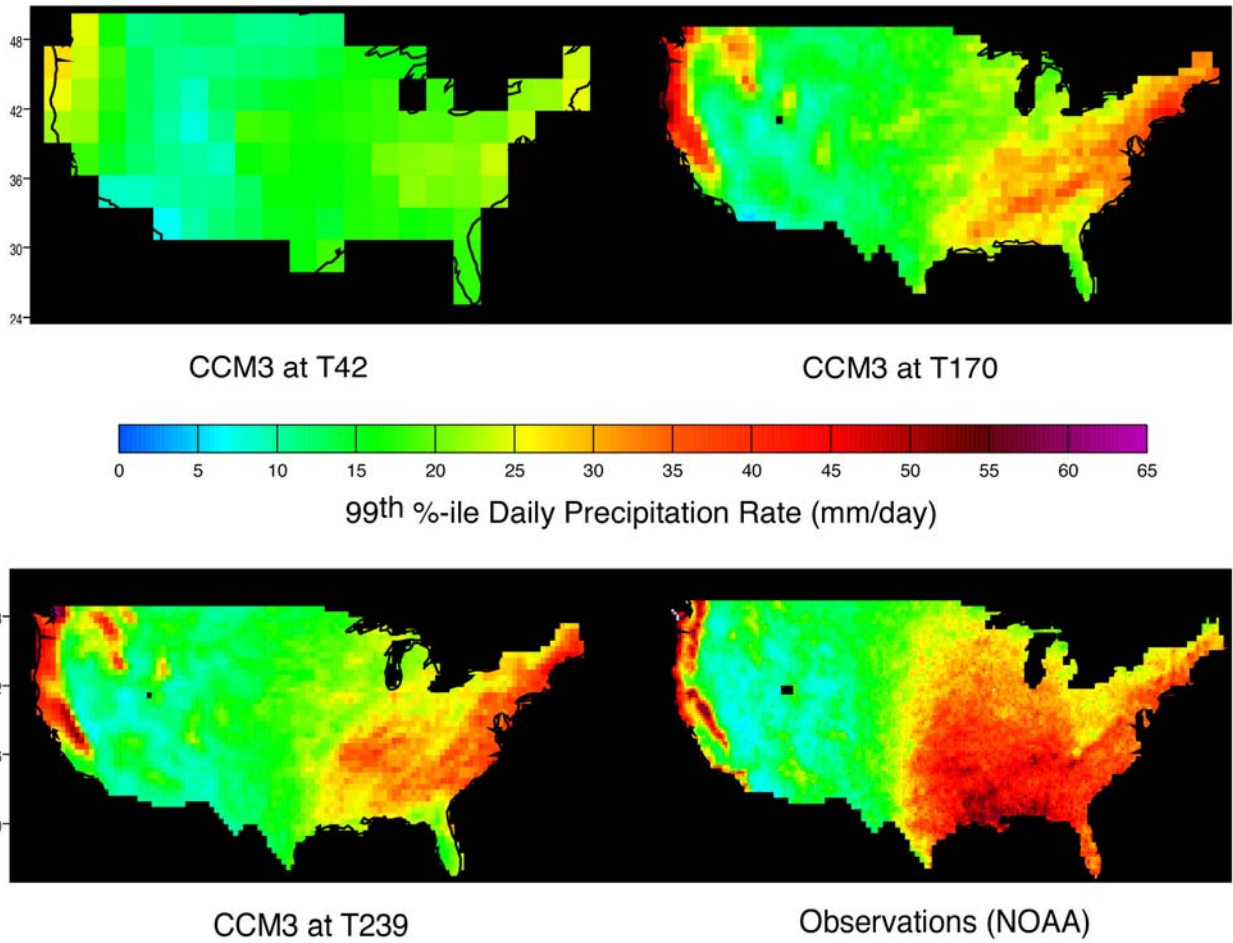
Figure A-6. Ninety-ninth percentile daily precipitation amounts, as simulated by the CCM3 global atmospheric model at truncations of T42 (~300 km resolution), T170 (~75 km resolution), and T239 (~50 km resolution), and in a NOAA observational data set. The 99th percentile daily precipitation amounts are one measure of the intensity of extreme precipitation events; this analysis shows that increasing spatial resolution greatly improves the ability to simulate these events. Iorio et al. (2004) discuss more thoroughly the effects of model resolution and sub-grid scale physics on the ability to simulate daily precipitation amounts in the continental United States.

CA SNOW DEPTH IN MARCH

Observation @ 5x5 km   CCM3 @ 50x50 km

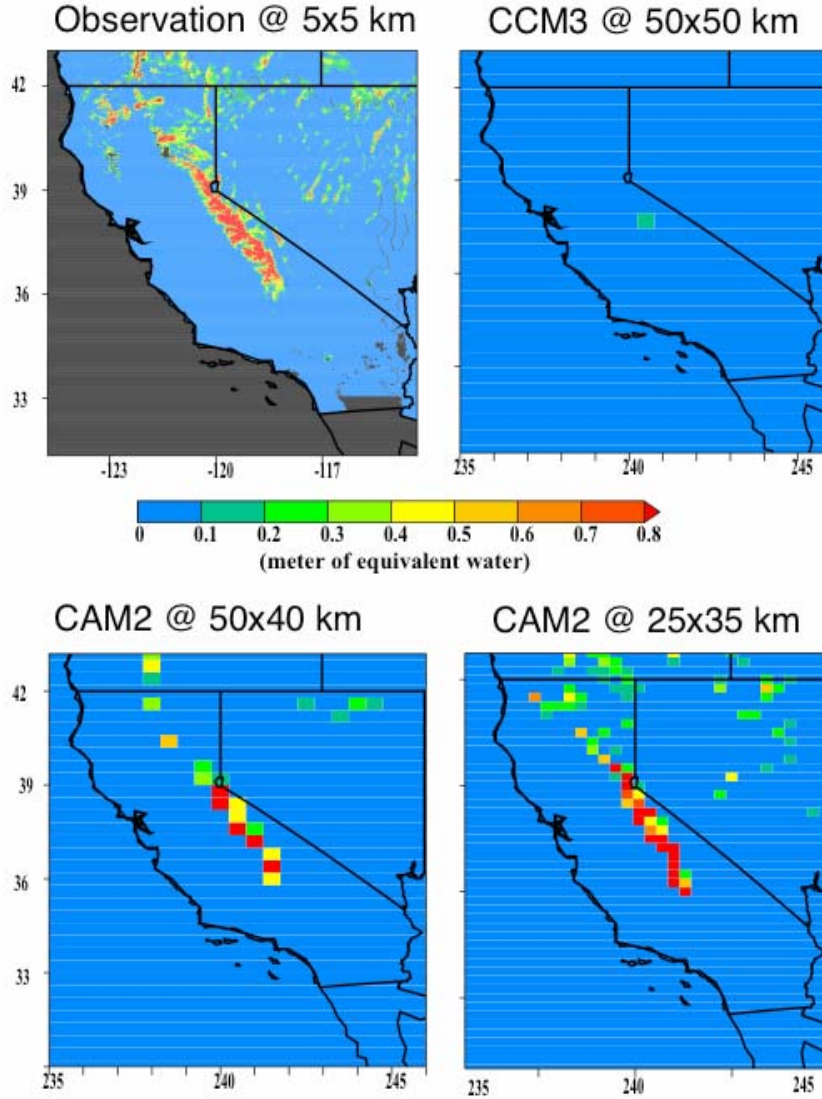(meter of equivalent water)

CAM2 @ 50x40 km   CAM2 @ 25x35 km

Figure A-7. Monthly mean water-equivalent snow depth for March. The top left panel shows NOHRSC observations; the top right panel shows results from the CCM3 global model at T239 truncation (~50 km resolution). The bottom panels show results of the NCAR CAM2 global model at 0.4 x 0.5 deg. resolution (left) and 0.25 x 0.35 deg. resolution (right; the latter simulation is ongoing). The CAM2 simulations use the finite-volume dynamical core, which allows better representation of orography than is possible with the Eulerian spectral dynamics used in the CCM3 simulation. This improves representation of snow by allowing surface elevations (and hence temperatures) in mountainous regions to be more realistic.